IED Dataset Generation: Analysis Across Theaters

Stephanie J. Wangeman

Abstract— Improvised Explosive Devices are the most dangerous threat to United States troops in Iraq and Afghanistan. Regardless of where troops deploy next, there is a strong likelihood of IED use in future conflicts. While great progress in analysis and exploitation of IED data has been made, the data is generally limited to historical attacks in current theaters of operation. The purpose of this analysis is to develop a tool that generates IED placement data based on statistically valid assumptions and distributions in any theater. These distributions can be used to train analysts prior to deployment to new environments. The research will generate data representing hypothetical IED attacks in a table structured like the CIDNE database. IED data from the CIDNE database will be used to develop distributions for each attack detailing IED location and type (type includes: command detonated, victim detonated, timer operated and radio controlled). Hypotheses will be tested using CIDNE data to determine some trends, including: seasonality, distance from road and location relative to terrain features. Results of these tests will shape the distributions used in the generation of synthetic datasets. The datasets will be generated based on road networks, terrain and population centers in Colorado, but shape files could be used from any region of the world. The results will generate various datasets that can be used for training analysts on IED attacks specific to any Operating Environment. This simulation will expose analysts to realistic data better preparing them for combat operations in their theater.

I. INTRODUCTION

Improvised explosive device strikes worldwide increased in 2012 despite a steep decline in Afghanistan [4]. While the IED threat initially caught international attention during the early stages of the War on Terror, it has become apparent that IED’s are a menace worldwide. IED’s offer the enemy the opportunity for maximum effectiveness with minimum commitment in terms of lives and resources.

Methods for combating the threat focus on detection and neutralization. Central to detection is analysis. Factors included in this analysis are seasonality, soil type, surrounding elevation, route traveled, enemy residence etc.. These factors control trigger type, charge amount, and materials used to create the IED. Charge amount and materials used affect damage implemented, while trigger type is determined by soil, signal and resources available.

Current analysis philosophy focuses on previous enemy IED placement to predict future trends in the form of spatial point analysis. While this philosophy is useful in current theaters, it does not focus on future theaters.

Proper prediction of IED type and placement allows for appropriate allocation of resources. Resources include metal detectors and signal jammers to name a few. This allocation promotes reduction in deaths due to IED’s, as it allows for IED detection. As it stands, a basis of previous enemy IED placement and patterns is necessary to determine resource allocation. The recent reduction in IED related deaths in Afghanistan reflects knowledge gained from current forms of IED analysis.

Fig. 1. A graph of ‘IED Casualties’ to ‘Year’ in Operation Enduring Freedom. It demonstrates the significance of IED related training, analysis and detection. Analysis leads to placement of correct resources, greatly reducing the IED threat. Adapted from [6].

A. Literature Review

Trend analysis highlights enemy location, material and trigger type. Trigger type corresponds to different soil types, elevation distributions and relative enemy location. There are four main triggers: radio, timer, command, and suicide. Radio triggered IED’s are detonated remotely, generally by an enemy who can see the situation unfolding but who cannot be picked out as unfriendly among the landscape. In Iraq and Afghanistan these IED’s are often detonated via cell phone. Timer operated IEDs are detonated via timer, therefore running the risk that they will not detonate at the appropriate time. Victim operated IED’s often make use of plates made of various materials surrounding charges which detonate under pressure. They are set off by a walking or driving, depending on the pressure required. These IED’s are buried just below the earth’s surface, often rendering them undetectable to sight unless the soldier has the opportunity to note the freshly buried signature.
turned soil or newly laid asphalt. Suicide IEDs are self-explanatory but are difficult to combat, as the enemy will not generally appear dangerous until they have detonated [1].

Current IED analysis focuses on trigger type, material and enemy. Trigger type offers a universal perspective, as a variety of triggers are generally available worldwide. Which materials are used and enemies detonate the IED’s are dependent on the resources available and enemy type in the region.

IED analysis allows researchers to identify different triggers in different locations based on human and geographical factors, promoting correct allocation of resources. This process of resource allocation promotes IED safety. For instance, it does not make logistical sense to send numerous signal jammers to a region which faces mostly victim operated IED’s [7].

When entering a new theater, it is difficult to ascertain where resources are best sent. The difficulty resides in the current focus of IED distribution analysis which relies on IED’s already detonated and patterns generated from analysis of these detonations. Current analysis fails to identify specific trends in the form of predictive models which may aid resource placement before gathering data resulting in an overall reduction of lives lost. The recent drop in deaths resulting from IED strikes in Afghanistan has been almost wholly attributable to resource distribution and training based on analysts’ predictions. This type of futuristic analysis in other theaters could help prevent loss of life [7].

B. Summary of Contributions

Current IED distribution analysis has done much to promote allocation of resources in Operations Iraqi Freedom and Enduring Freedom. With IED incidents increasing worldwide, however, the threat will remain after current operations cease. Resultantly, it is important to look for a distribution analysis system which relies on key trends already discovered in current theaters to train analysts to recognize initial resource allocations which would prevent casualties earlier in the conflict. Current trends already analyzed based on soil type, elevation, enemy location etc. can inform a model which would allow analysts entering new theaters to train to the new theater and its various terrain features affecting IED distribution to determine an appropriate initial allocation of resources before real data is available.

II. PROBLEM DEFINITION

IED analysis focuses on detonations and resource allocated in current theaters. While this analysis has done much to reduce the IED threat in Afghanistan, it has done little to curtail the increasing threat worldwide. Proper resource allocation is the main tool for combating IEDs. For instance, if, in an area in which most IEDs are radio detonated, and a simple signal jammer could prevent the majority of IED bursts, and troops are given metal detectors, casualties will skyrocket. If, instead, analysts are able to analyze area soil types and materials available, before entering the theater, perhaps ascertaining that cell phone towers are not present in a particular area, so soldiers should have metal detectors and should also note that most IED’s in this mountainous region are placed near good ambush positions, the soldier has been given significant information which could save his life even before friendly forces have entered the theater.

Also key to IED dataset generation is training for analysts. If an analyst can build potential detonation distributions before entering the theater, she can train to see trends in these distributions, giving her valuable experience in instinctively identifying potential distributions in any theater. As intelligence assets are still being set up in new theaters, analysts can use their training to affect ground operations without analyzing strikes which have already been made.

A. Improved Analysis

A new system for IED analysis would not only allow for improved training experiences, it could lay the foundation for a more intuitive analysis system. Current systems are friendly to those who are good with computer programming and excel. A system with foundations in R and applications which are made to be user friendly (with an interface with Google Earth for instance) would simplify usage and place focus on results instead of trend generation.

Also, a program which generates datasets would not have the same constraints as a system classified as secret. It would allow analysts who do not yet have their secret classification to train on a system similar to what they would use in theater. This model has the potential to universalize IED analysis. Systems which currently rely on excel and exist only on secret networks could be utilized and understood by those who are not excel experts. This would provide an interface which would allow for better communication with commanders and logistics officers who determine resource allocation.

B. Potential Model Value

The goal of this model is to allow analysts to train in and to offer resource allocation suggestions for new theaters. Existing on both the secret and unclassified side of the network, the model could, in its advanced stages, universalize and simplify methods used for IED placement analysis, allowing friendly forces to get a jump on potential enemy action and to predict trends which, before entering new theaters, would not otherwise be apparent. These trends, at least initially, will rely on factors present in every theater such as: road distribution, soil type, surrounding population (metropolitan or rural area), surrounding elevation and routes commonly traveled.

The model’s current foundations in R allow for the development of a user interface which is friendly, can be
used on web platforms and in combination with powerful applications like Google Earth. R’s numerous packages also allow those who are good with programming in R to easily manipulate the data and to develop new systems for analyzing new trends which could then undergo similar processes, becoming more user-friendly and developing systems with interfaces in other programs to continue to universalize and develop IED analysis.

View this model as a starting point which offers an opportunity to stay one step ahead in any theater and which also serves as an algorithm which can be used to continue to develop a flexible system used to analyze IED distributions in any theater using factors common to those theaters. The system can then be manipulated to fit current theaters and used both on the unclassified and secret sides of the network. Presence of the model on both sides of the network promotes greater unity of focus and better training for those who have not yet entered current theaters.

**D. Requirements**

When building a training system for analysts, it is important to acknowledge which factors apply to the majority of theaters, and which may only be present in current theaters. For instance, in some theaters IED distribution may not correlate to where the enemy resides, but instead where friendly units operate. Perhaps enemy forces are very mobilized and travel far from homes and bases.

Another important factor is whether the data appears intuitively correct. Shapefiles can result in a convoluted appearance of data. It will be important to display the data in layers which can be manipulated both to allow analysts to understand how different factors affect IED placement and to promote clarity in viewing the final product.

The analysts must also have the ability to easily manipulate the system. Analysts should be able to change the distributions used to generate the data as current theaters develop to allow the model to best reflect current technologies and trends. Current systems for analysis are friendly to computer programmers and to those with a talent for excel. Future systems should be intuitive and allow analysts to quickly see trends with little effort. A visual interface is key.

### III. SYSTEM DESIGN

The model was developed using Rstudio and shapefiles from the TIGER/Line® database for Weld County, Colorado. The system currently only runs in R, but there is potential to expand its capabilities to interface with the R package, RShiny, which produces an interactive web-based interface or Google Earth, which can be integrated with R based data generation.

This model is about creating a flexible system which will allow for IED generation in any theater. System generation occurs in four steps: 1) choosing the theater 2) determining factors for analysis 3) generating distributions based on the factors and 4) modeling in a visual setting.

Choosing the theater is self-explanatory, which leads to a discussion of factors for analysis and how to display those factors, followed by distribution generation and modeling in a visual setting.

**A. Mapping**

Displaying IED detonations in a visual interface presents challenges in terms of the data to be displayed, the numerous factors which play into IED placement, and whether the model is visually intuitive. These considerations make it apparent that focusing on a nation or a theater as a whole is not feasible. There are numerous layers required to present an accurate picture of the main factors of IED placement. Layers include: soil type, elevation relative to potential target, bodies of water, seasonality, distribution of road, and area population.

![Fig. 2. An example of ‘Ideal Points for IED’s.’ It demonstrates some key considerations in IED distribution analysis to include soil type, road shape and surrounding elevation. Adapted from [1].](image)

It is apparent that these layers cannot be displayed on a single map. In fact, analysis of a particular theater may rely on the ability to zoom in on a small operational area. Regardless of the number of IED’s plotted, the confusion created by the numerous layers would render the visual display useless to analysts.

**B. Value of Data in R**

Shapefiles utilized in R offer integration between physical features and distributions used to plot spatial points. Rather
than creating a distribution as an overlay and plotting randomly generated points, R allows the user to interact with the shapefile, plotting points on the roads themselves, rather than as a distribution overlay, as might be generated by using Google Earth based applications.

Generating the distributions is simple enough and involves basic statistical analysis based on trends in data. Plotting the distributions on a visual interface is what makes shapefiles so vital to the process.

R allows the user to generate random distributions based on the shapefiles present and the relationship between the features in the shapefile. For instance, R allows the user to build a distribution based not only on road location, but on interpolation between points used to generate roads in the shapefiles, massing points on curves and intersections as occurs in theater and as is demonstrated in Figure 2.

R also allows the user to write shapefiles of spatial points generated. This allows the analyst to change various factors which affect IED distribution and save both the coding and the visual accompaniment to share with other analysts and to utilize in other programs which facilitate shapefiles.

C. Manipulating the Model

As it stands, the model is in its rudimentary stages of development. Many questions remain. What layers will be present in the model? Considerations of social networking have aided the current intelligence community in tracking targets and enemy populations. Could social networking be included in the model?

It is clear that seasonality, soil type, etc. play a big role. What is not clear is how much value a model which takes into account social networking and other factors that may be highly specific to a particular theater could have.

Also up for debate is whether or not R presents the best interface for the model. While it is clear that R offers intimate interaction between distribution development and plotting spatial points, not to mention shapefile generation, it may not offer the most intuitive user interface. While numerous packages, such as RShiny, present the opportunity for a web interface which can be manipulated, it is not clear that a map, such as what is presented in Figure 3 could be useful to analysts. The convoluted nature of the shapefile display presents concerns.

D. The Importance of the Process

Regardless of concerns over R, what layers should be included, etc., at this early stage of development, focus must remain on establishing a uniform process and system which analysts can use for training or for combat to develop a better understanding of the threat.

As it stands, IED attacks are a worldwide phenomena which, with a recent increase in 2012, are showing no signs of stopping. Training and data generation in potential theaters will allow analysts to promote the correct allocation of resources, preventing deaths early on and to combat the power of IED’s.

IED’s are unique insofar as they allow the enemy to inflict mass casualty with little loss. Key trends in data have been detected and present the opportunity to train analysts via useful and uniform interfaces which allow for quicker data integration and analysis.

IV SYSTEM TESTING

A. Correctness According to Current Analysts

“Maps reveal patterns, and patterns provide insight. Analytic maps can be used to identify the patterns in the environment and explain them to others [2].” While the current model has not yet reached the testing stage, a good
model would appear correct to analysts with experience in the field. Distributions can be validated, models created, graphs displayed, but all of this is useless if analysts do not appreciate the model or do not feel it displays data which is similar to that which they see in the real world.

B. Flexibility

Also, analysts in the theater currently use programs to automatically generate various graphs displaying overall IED trends, potential for increased threat and even heat maps for potential future attacks based on particular routes. In developing an intuitive map-based model, considerations of current data generated and how that data is communicated to commanders cannot be lost. It is important in the context of these displays, however, to determine which are most useful as part of a uniform model which functions across theaters.

Analysts currently in the field should also be able to manipulate the model in intuitive and practical ways. Access to the programming and distributions used is necessary. Also, the ability to add more layers, such as social networking, will allow the model to function both as a training tool and as a tool in theater.

C. Patterns and Accessibility

Analysts in training face the difficulty of not observing real IED data until deployed. Artificial IED data, generated from distributions found in real data and displayed in different environments grants analysts the opportunity to work with the data and to notice patterns analysts in the field are already intimately familiar with. Part of the accessibility of data will rely on opening the conversation with current analysts. A good model will not only present data which is visually accessible, but data which highlights those patterns which have allowed analysts to successfully reduce the IED threat. “Effective analysis and dissemination of IED intelligence allows for the development and implementation of tactics to defeat the latest threats, identify the perpetrators of IED attacks and bring them to justice [7].”

Factors to consider when developing such a model include realistic display of imagery. A road should look like a road, a building like a building and a mountain like a mountain. This display makes the model more intuitive.

IV. NEXT STEPS

As it stands, the current model allows users to utilize distributions to generate spatial points onto imagery generated via shapefiles in R. The user can utilize any layer for which there are shapefiles available in the region, and any distribution which matches current IED data. It offers an interesting glance into what is possible with a system for generating IED’s in any theater, and the potential life-saving observations which can be made about new theaters, not to mention the valuable training gained by those who have not yet entered the current theater.

The model encounters difficulties in terms of friendly user interface, mapping onto a realistic visual surface and capability for displaying numerous layers at once. These considerations call into question the validity of the current model in serious ways. This model is intended to serve as the foundation for further work. It begs the question: how can we make this happen? It brings to light just how valuable this sort of training can be.

As focus in war shifts from larger targets inflicting a lot of damage, to smaller targets inflicting proportionally larger damage in terms of IED’s and even cyber warfare, the concepts behind this model offer a glimpse into a unified counter-front. More and more of war fighting focuses on enemy analysis, an understanding of social networking and, fundamentally, recognition of patterns in any environment. This model offers two key opportunities for analysts. The first is that it allows analysts to train in any theater. The second is that it allows analysts who have the proper shapefiles to generate potential IED attacks in their next theater. Information gained will aid commanders in anticipating where to allocate their resources to combat the IED threat.

VI. CONCLUSION

IED’s are currently the most dangerous threat to troops in Iraq and Afghanistan. As the nation expands into new theaters, the IED threat shows no sign of diminishing. IED’s are currently the number one killer in Operation Enduring Freedom. Their deadliness highlights the need for a uniform system which can be used by analysts and trainees alike to better understand the IED threat and to predict the patterns present in upcoming theaters. An intuitive system which allows for better prediction will give commanders the information they need to allocate resources more quickly, saving lives. While it is still in its rudimentary stages, this model offers the framework for a system which will be used to develop a uniform system used by analysts across theaters. Which specific factors should be included in the system will be dictated by current analysts and future theaters. Whatever the factors, it is clear that such a system will give commanders information necessary to train and aid their troops in combating this threat.

ACKNOWLEDGMENTS

Gratitude goes to LTC Stephen J. Henderson who set up the initial model based on Oregon which would lay the foundation for model development in other theaters. Thank you also to MAJ Michael W. Tilton whose advice was invaluable and discussions about the relevance of the overall system and process added great value to the work. Were it not for these advisors, the project in its current scope would not have been possible.
REFERENCES


